given situation when one or more (other) words related the situation are known. Thus, the system may change its operation based on observed facts, by using previous experience stored in the experience matrix of the system.

[0029] One or more prediction words may be determined from query word or words, by using the data stored in the experience matrix. The system may be subsequently controlled based on the prediction words. The query words may be words, which describe the current situation. The prediction words are words, which are likely to describe the current situation. Co-occurrence data may be stored as vectors of an experience matrix EX**1**. The co-occurrence data may be subsequently used for making a prediction. A vocabulary VOC**1** may be needed to store co-occurrence data in the experience matrix EX**1** and to utilize co-occurrence data stored in the experience matrix EX**1**. In an embodiment, the vocabulary VOC**1** may be used also in connection with another matrix, called the basic matrix RM**1**.

[0030] Referring to FIG. **1***a*, a vocabulary VOC**1** may comprise a group of different words $w_1, w_2, \ldots w_n$. The vocabulary VOC**1** may be a hash table, which contains pointers to sparse vectors of a basic matrix RM**1**. The sparse vectors may also be called as random index vectors (RI vectors). Thus, each word $w_1, w_2, \ldots w_n$ of the vocabulary VOC**1** may be associated with a basic sparse vector $a_1, a_2, \ldots a_n$. For example, a word w_k (e.g. "meeting") may be associated with a basic sparse vector a_k. The basic sparse vectors $a_1, a_2, \ldots a_k, \ldots a_n$ may be stored e.g. as rows of the basic matrix RM**1**.

[0031] The vocabulary VOC**1** may be a hash table, which indicates the location (e.g. row k) of the sparse vector in the basic sparse matrix RM**1**, by using the word (e.g. w_k) as the key.

[0032] Each basic sparse vector of the matrix RM**1** may represent a word. For example, the basic vector a_k may be interpreted to represent the word w_k in sparse vector format. Each basic sparse vector a_k consists of elements $R_{1,k}, R_{2,k}, \ldots, R_{ik}, \ldots, R_{m,k}$. In an embodiment, the basic sparse vectors $a_1, a_2, \ldots a_k, \ldots a_n$ of the matrix RM**1** may be unique and different. Each row of the basic matrix RM**1** may be a unique basic sparse vector associated with a different word. Each basic sparse vector a_k may have a high number of zero elements and only a few non-zero elements. For example, a basic sparse vector a_k may have e.g. 10000 elements R wherein twenty elements may be non-zero and 9980 elements may be zero.

[0033] In an embodiment, the sum of all elements of the basic sparse vector a_k may be equal to zero. This may minimize memory consumption, may simplify mathematical operations and/or may increase data processing speed. In particular, 50% of the non-zero elements may be equal to −1 (minus one), and 50% of the non-zero elements may be equal to 1 (one). In other words, the value of an element may be −1, 0 or 1, and the basic sparse vector a_k may be a ternary vector.

[0034] Each vector $a_1, a_2, \ldots a_k, \ldots a_n$ may represented by a point in a multi-dimensional space. More precisely, each vector $a_1, a_2, \ldots a_k, \ldots a_n$ may represented by a different end point in the same multi-dimensional space when the starting point of each vector is located at the same point (e.g. origin). The number m of elements R of the basic sparse vector a_k may be e.g. in the range of 100 to 10^6. The number m_{nz} of non-zero elements R of the basic sparse vector a_k may be in the range of 0.1% to 10% of the number m, said number m_{nz} of non-zero elements also being in the range of 4 to 10^3. Increasing the total number m and/or the number m_{nz} of non-zero elements may allow using a larger vocabulary VOC**1** and/or may provide more reliable predictions. However, increasing the number m and/or m_{nz} may also require more memory space and more data processing power.

[0035] R_{ik} denotes an element of the basic matrix RM**1** belonging to the column and to the k^{th} row. In an embodiment, the vectors may be ternary and the value of an individual element e_{ik} may be one of the following −1, 0, or 1. The number of negative non-zero elements R may be equal to the number of positive non-zero elements R, the values of the non-zero elements R being integers. This likely to maximize data processing speed, and to minimize the use of memory. However, this is not necessary. For example, a basic sparse vector a_k may have sixteen elements R having a value −0.5 and four elements R having a value 2.

[0036] When the vectors $a_1, a_2, \ldots a_k, \ldots a_n$ of the matrix RM**1** are compared with each other, the positions of the non-zero elements in a basic sparse vector a_k and the values of the non-zero elements of the basic sparse vector a_k may be randomly distributed. The basic sparse vectors may also be called as random index vectors. This randomness may ensure with a high probability that the points representing the basic sparse vector in the multi-dimensional space are not too close to each other. If two points representing two different words would be too close to each other, this might lead to erroneous predictions during subsequent processing. When the positions and the values are randomly distributed, this may also ensure with a high probability that each word of the vocabulary VOC**1** is associated with a unique and different basic sparse vector. When the indices are random, it is highly probable that the elementary sparse vectors associated with two different words are orthogonal or nearly orthogonal. Thus, the dot product of said elementary sparse vectors is equal to zero at a high probability. This pseudo-orthogonality of the words may preserve the unique identity of each word or event or unique occurrence stored in the matrix, even when they are represented by the sparse vectors. This pseudo-orthogonality of the words may preserve the unique identity of each word, event or unique occurrence stored in the experience matrix EX**1** when the words, events and/or occurrences are represented as combinations of the sparse vectors in the experience matrix EX**1**. The words of the vocabulary may be arranged e.g. in alphabetical order.

[0037] FIG. **1***b* shows several sets of words, which may be called e.g. as "bags". A bag comprises two or more different words related to each other. The number of words in a bag may be e.g. in the range of 2 to 1000. Each bag could also be called as a "document".

[0038] The bags represent co-occurrences of words. The bags may be used to educate a system. Information about co-occurrences of the words may be stored in a matrix EX**1** (FIGS. **1***c*, **4***b*).

[0039] The words of a bag may e.g. describe a situation and/or a system used in said situation. In particular, the words may represent observed facts.

[0040] For example, a first bag BAG**1** may comprise the words "work", "morning" and "mail". The word "work" may be e.g. determined e.g. based on a marking in the calendar of the system. The word "morning" may be provided e.g. by a clock of the system. The word "mail" may be related to the words "work" and "morning" e.g. because the system has detected that a user has received and/or sent written messages in a situation described by the words "work" and "morning".